

NASA TECHNICAL NOTE



NASA TN D-8251

NASA TN D-8251

LOAN COPY: RE  
AEWL TECHNICAL  
KIRTLAND AFB



THE DAVIDON-FLETCHER-POWELL PENALTY  
FUNCTION METHOD: A GENERALIZED  
ITERATIVE TECHNIQUE FOR SOLVING  
PARAMETER OPTIMIZATION PROBLEMS

*Ivan L. Johnson, Jr.*

*Lyndon B. Johnson Space Center  
Houston, Texas 77058*



NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • MAY 1976



0133994

1. Report No. TN D-8251		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle <b>THE DAVIDON-FLETCHER-POWELL PENALTY FUNCTION METHOD: A GENERALIZED ITERATIVE TECHNIQUE FOR SOLVING PARAMETER OPTIMIZATION PROBLEMS</b>		5. Report Date May 1976		6. Performing Organization Code JSC-09682	
7. Author(s) Ivan L. Johnson, Jr.		8. Performing Organization Report No. S-447		10. Work Unit No. 986-16-00-00-72	
9. Performing Organization Name and Address Lyndon B. Johnson Space Center Houston, Texas 77058		11. Contract or Grant No.		13. Type of Report and Period Covered Technical Note	
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code			
15. Supplementary Notes					
16. Abstract The Fletcher-Powell version of the Davidon variable metric unconstrained minimization technique is described. Equations that have been used successfully with the Davidon-Fletcher-Powell penalty function technique for solving constrained minimization problems and the advantages and disadvantages of using them are discussed. The experience gained in the behavior of the method while iterating is also related.					
17. Key Words (Suggested by Author(s)) Constraints Nonlinear programing Optimization Minimization Trajectory optimization Parameter optimization Constrained minimization Steepest descent method			18. Distribution Statement STAR Subject Category: 64 (Numerical Analysis)		
19. Security Classif. (of this report) Unclassified		20. Security Classif. (of this page) Unclassified		21. No. of Pages 24	
				22. Price* \$3.50	



## CONTENTS

Section	Page
SUMMARY . . . . .	1
INTRODUCTION . . . . .	1
SYMBOLS . . . . .	3
GENERAL DESCRIPTION . . . . .	4
CONSTITUENT ELEMENTS . . . . .	6
DFP Method . . . . .	6
Penalty Function Method . . . . .	13
CHARACTERISTICS . . . . .	16
CONCLUDING REMARKS . . . . .	19
REFERENCES . . . . .	20

# THE DAVIDON-FLETCHER-POWELL PENALTY FUNCTION METHOD: A GENERALIZED ITERATIVE TECHNIQUE FOR SOLVING PARAMETER OPTIMIZATION PROBLEMS

By Ivan L. Johnson, Jr.  
Lyndon B. Johnson Space Center

## SUMMARY

The Davidon-Fletcher-Powell penalty function method is a technique that has been used successfully to solve constrained minimization problems. The method was devised by combining an exterior penalty function with a performance function for solving constrained minimization problems.

Constrained minimization problems include parameter optimization, optimal control, boundary values, and iterative problems. In addition to relating the experience obtained while observing the technique, this report addresses the successful application of the Davidon-Fletcher-Powell penalty function technique in combination with double-precision arithmetic for solving atmospheric and exoatmospheric flight optimization problems.

## INTRODUCTION

One of the most significant advances in numerical optimization in the last 5 years has been the use of finite difference arithmetic to determine partial derivatives for numerical gradient optimization iterative methods. This single achievement has substantially increased the level of problem statement complexity that can be subjected to numerical optimization methods. The purpose of this report is to describe one such iterative method, the Fletcher-Powell version of the Davidon variable metric unconstrained minimization technique, using finite difference gradient information together with an exterior penalty function technique for solving constrained minimization problems.

The Davidon-Fletcher-Powell (DFP) method was combined with an exterior penalty function method to provide the capability for solving constrained minimization problems (ref. 1). Subsequently, a computer program of the DFP penalty function method was developed and has been used continuously to solve many types of parameter optimization problems.

One of the first types of problems solved was an impulsive orbit transfer optimization problem (ref. 2) wherein the components of the velocity impulses and the transfer angles were defined as variables or parameters. Other types of problems solved were in the areas of optimal nuclear reactor design, optimal aircraft control

systems design (ref. 3), optimal feature selection (Earth resources), curve fitting data, solving systems of equations, optimal electronics problems (radar), and atmospheric and exoatmospheric flight optimization<sup>1 to 5</sup> (refs. 4 to 7). With the exception of the last two problems, the DFP penalty function method was used with closed-formed gradients, or partial derivatives, on all these problems. The optimal electronics problem involving a simplified design of a radar system was solved by using numerical, central, and forward difference gradients computations. Before this problem was solved, the accuracy of the DFP method in using numerical partial derivatives was questionable, even though the optimal electronics problem was a simple problem with only seven parameters. The atmospheric and exoatmospheric flight optimization problems were also solved by using numerical gradients, but some of these problems involved as many as 60 parameters and included trajectories integrated over thousands of seconds. However, the DFP penalty function method apparently worked almost as well with numerical gradients as with closed-form gradients.

In solving atmospheric flight optimization problems, many techniques were used to satisfy state and control inequality constraints. One technique used, whenever possible, was that of solving for the control explicitly from the constraint equation when the constraint was violated. However, a different technique, an error function method similar to the integral penalty function method, was used when this was not possible. This method provides a better way of shaping a control and satisfying state and control inequality constraints<sup>6</sup> (ref. 8).

A tried and proven computer program of the DFP method was devised by the Mathematical Analysis Section of the Software Development Branch at the NASA Lyndon B. Johnson Space Center. With this program, the user can select either closed-form gradients, central difference gradients, or forward difference gradients. The one-dimensional search that is used by the DFP method offers three

---

<sup>1</sup>I. L. Johnson and J. L. Kamm, "Performance Analysis of 049A/SRM Shuttle for Maximum GLOW," JSC IN 72-FM-202, Sept. 1972.

<sup>2</sup>J. L. Kamm and I. L. Johnson, "An Optimal Solid Rocket Motor Thrust Profile to Minimize 040A Shuttle GLOW," JSC IN 72-FM-159, June 1972.

<sup>3</sup>I. L. Johnson, "A Minimum Sized Space Shuttle Configuration," JSC IN 72-FM-161, June 1972.

<sup>4</sup>I. L. Johnson, "Minimum Fuel Deorbit Using PEACE Program," JSC IN 72-FM-136, May 1972.

<sup>5</sup>I. L. Johnson, "Sizing a Siamese Lifting Body Triamese Tank Ascent Configuration," JSC IN 72-FM-259, Nov. 1972.

<sup>6</sup>I. L. Johnson, "The Error Function Method for Handling Inequality Constraints," JSC IN 73-FM-114, July 1973.

options: cubic fit using the slope at the initial point, cubic fit without the slope, and the golden section method<sup>7</sup> (refs. 9 and 10<sup>8</sup>). Interfacing this program with another program is both simple and flexible. A version of this program is in the space vehicle dynamics simulation (SVDS) program and is currently in the checkout phase. The DFP penalty function program, which can be obtained from the author, is sometimes referred to as the parameter optimization program, the accelerated gradient program, or the Davidon program.

## SYMBOLS

$a_0, a_1, a_2, a_3$	polynomial coefficients
$a_i$	polynomial coefficient
$f(x)$	general performance function
$g(x), h(x)$	vectors of point constraint functions
$H$	positive-definite symmetric matrix
$I$	identity matrix
$K_1, K_2$	diagonal weight matrices
$n$	number of variables or parameters
$p$	penalty and quadratic function
$p(x)$	penalty and quadratic function of $x$
$p_x$	gradient of $p(x)$
$p_{xx}$	matrix of second partials with respect to $x$
$s_i$	direction of search on $i$ -th iteration
$U(h)$	diagonal matrix of unit step functions
$U(h_j)$	diagonal elements of $U(h)$
$v$	speed

---

<sup>7</sup>I. L. Johnson and G. E. Meyers, "One-Dimensional Minimization Using Search by Golden Section and Cubic Fit Methods," JSC IN 67-FM-172, Nov. 1967.

<sup>8</sup>The correction to this paper can be obtained from the author.

$x$	$n$ -dimensional parameter vector
$y_i$	difference in gradient vectors for $i$ -th iteration
$\gamma$	one-dimensional search parameter
$\gamma_i$	value of $\gamma$ on the $i$ -th iteration for the minimum of $p(x_i + \gamma s_i)$
$\varepsilon$	small positive number
$\lambda$	Lagrange multiplier

#### Subscripts:

$f$	final value
$LB$	lower bound
$x$	partial derivative with respect to vector $x$

#### Operators:

$T$	transpose
$-1$	inverse
$-$	at minimum and desired value
$'$	ordinary derivative

## GENERAL DESCRIPTION

The constrained minimization problem, or nonlinear programming problem, is defined as finding the  $n$  dimensional parameter vector  $x$  that locally minimizes the general performance function  $f(x)$  and that satisfies the vectors of point constraints  $g(x) = 0$  and  $h(x) \leq 0$ . A popular method for solving this problem is the gradient projection method. This method was initiated by Rosen (refs. 11 and 12) and later improved by Kelley and Speyer (ref. 13) and others. It requires larger amounts of computer storage than others, and it requires the computation of an inverse of a matrix. Because of linearization assumptions, this method may also require small step sizes. The user must be knowledgeable in selecting numerous constants associated with the gradient projection method because the transition between performance function minimization and constraint satisfaction must be properly balanced.



The penalty function method provides an alternate approach for solving the constrained minimization problem by solving an unconstrained minimization problem. The performance function together with a term that "penalizes" it for violations in the constraints is minimized. Penalty functions are divided into two classes: interior and exterior.

The interior penalty function method by Fiacco and McCormick (refs. 14 and 15), which was popularized by Brusch (ref. 16) and others, is probably the best interior method currently known. Although performance using this method is good, an initial parameter vector that satisfies the point inequality constraints is needed.

An exterior penalty function method (refs. 1, 17, and 18) allows initial parameter vectors to cause violations in inequality constraints. This method has low computer storage requirements and a smooth, continuous transition between performance function minimization and constraint satisfaction. The exterior method, as well as the interior method, is minimized by an unconstrained minimization method.

A survey of the literature reveals many promising unconstrained minimization methods derived by using various assumptions. Probably the most basic method is sectioning. This method simply produces a direction of search, by changing only one parameter per iteration. Although sectioning does not require the computation of costly gradient vectors, it requires a very large number of function computations and small steps.

The method of steepest descents (ref. 1), with a direction of search essentially in the direction of the negative gradient vector, is based on linear assumptions and thus requires small steps. Convergence is slow with the steepest descent method, but it is faster than sectioning. The pattern search can be used to speed up this method and others by periodically forming a direction of search with a vector that is the resultant of  $k$  steps. Linear assumptions are still a disadvantage in using the method of steepest descents.

Several methods derived by assuming that the function to be minimized is quadratic take steps in directions that are conjugate with respect to the matrix of second partial derivatives. For a quadratic function of  $n$  variables, if the steps are found by performing successive, one-dimensional minimizations in these conjugate directions, then convergence to the minimum is obtained within roundoff in, at most,  $n$  steps (ref. 19). When applied to a nonquadratic function, the function to be minimized appears to be quadratic as the minimum is approached. (The second-order terms in the Taylor series expansion of the function dominate because the first-order terms are vanishing.) These methods accelerate convergence, whereas methods based on linear assumptions, such as steepest descent, tend to slow it down.

Two of these methods, the conjugate gradient method of Fletcher and Reeves (ref. 19) and a method devised by W. C. Davidon (ref. 20) that was later modified by Fletcher and Powell (ref. 21), are currently the most popular. Compared to Davidon's method, the conjugate gradient method requires less computer storage and much simpler computation. In reference 22, however, Davidon's method featured much faster convergence for nonquadratic functions.

Because of the advantages, the Davidon-Fletcher-Powell method is used as the unconstrained minimization method in conjunction with an exterior penalty function for solving the highly nonlinear constrained minimization problems defined by optimization problems.

## CONSTITUENT ELEMENTS

### DFP Method

The DFP method is one of the most powerful iterative methods known for minimizing a general unconstrained function of  $n$  variables or parameters. This iterative method was derived by W. C. Davidon (ref. 20) and later modified and clarified by R. Fletcher and M. J. D. Powell (ref. 21). It is one of many derived by considering that the function to be minimized is a quadratic function.

Overall procedure.— Iterative minimization methods determine the position of a minimum  $\bar{x}$  of a function as the limit of a sequence of successive positions  $x_0, x_1, \dots$ . These positions are obtained by minimizing the function along a line through the previous position in a specified direction of search  $s_i$  obtained by some specified rationale. For a quadratic function  $p(x)$  with a regular minimum located at the position  $\bar{x}$ ,

$$p(x) = p(\bar{x}) + \frac{1}{2}(x - \bar{x})^T p_{xx}(x - \bar{x}) \quad (1)$$

In reference 19, it was shown that if a sequence of nonzero directions of search  $s_0, s_1, s_2, \dots$  is  $p_{xx}$  orthogonal or conjugate, that is,

$$\begin{aligned} s_i^T p_{xx} s_j &= 0 \text{ for } i \neq j \\ &> 0 \text{ otherwise} \end{aligned} \quad (2)$$

and if at the end of each search  $p(x_i + \gamma s_i)$  is minimized with respect to the one-dimensional search parameter  $\gamma$ , that is,

$$s_i^T p_x(x_{i+1}) = 0 \quad (3)$$

then the minimum of  $p$  is obtained in, at most,  $n$  iterations. An iterative minimization method that minimizes a quadratic function in a finite number of iterations, usually  $n$ , is said to have quadratic convergence (ref. 23).

For a nonquadratic function  $p(x)$ , a Taylor series expansion about the position of the minimum  $\bar{x}$  is given by

$$p(x) = p(\bar{x}) + \frac{1}{2}(x - \bar{x})^T p_{xx}(\bar{x})(x - \bar{x}) + \text{higher order terms} \quad (4)$$

From equation (4),  $p(x)$  behaves like a quadratic function near the minimum for all  $x$  in a small neighborhood about  $\bar{x}$ :  $\|x - \bar{x}\| < \varepsilon$ , where  $\varepsilon$  is a sufficiently small positive number. For values of  $x$  in this neighborhood, quadratically convergent, iterative methods will accelerate convergence to the minimum of the general function.

If given a position  $x_0$  for a quadratic function  $p$ , the position of its minimum  $\bar{x}$  can be found in one iteration by  $\bar{x} = x_0 + s_0$  where

$$s_0 = -p_{xx}^{-1} p_x(x_0) \quad (5)$$

and  $p_{xx}^{-1}$  is assumed to exist.

For a general function  $p$  beginning at the point  $x_0$ , the position of the minimum  $\bar{x}$  is found as the limit of the sequence  $x_0, x_1, \dots, \bar{x}$ ; each  $x_j$  is obtained by applying  $x_{i+1} = x_i + \gamma_i s_i$  where

$$s_i = -p_{xx}(x_i)^{-1} p_x(x_i) \quad (6)$$

Equations (5) and (6), sometimes referred to as Newton's method, require the computation of  $p_{xx}$  and its inverse. For a general function, computations of  $p_{xx}^{-1}$  would be performed at each point  $x_i$ . This procedure often would require excessive time because  $p_{xx}$  would probably be obtained by a numerical method rather than in closed form. Also,  $p_{xx}$  must be inverted, but this requirement may not be achievable because there is no guarantee that  $p_{xx}$  is nonsingular for all points  $x$ .

Instead of computing  $p_{xx}^{-1}$ , a positive-definite symmetric matrix  $H$  is computed iteratively in the DFP method such that a direction of search is given by

$$s_i = -H_i p_x(x_i) \quad (7)$$

For a quadratic function,  $H$  converges to  $p_{xx}^{-1}$ ; and for a general function,  $H$  tends to  $p_{xx}^{-1}$  evaluated at the minimum.

The DFP method begins with an initial point  $x_0$ , an initial matrix  $H_0$  (sometimes selected as the identity matrix  $I$ ),  $p(x_0)$ , and the gradient vector  $p_x(x_0)$ .

During the  $i$ -th iteration, the direction of search is computed from equation (7), and  $p(x_i + \gamma s_i)$  is minimized as a function of  $\gamma$ , a positive scalar step-size parameter. The position of the minimum is defined as

$$x_{i+1} = x_i + \gamma_i s_i \quad (8)$$

where  $\gamma_i$  is the value of  $\gamma$  on the  $i$ -th iteration for the minimum of  $p(x_i + \gamma s_i)$  as a function of  $\gamma$ . At the new position  $x_{i+1}$ , the gradient  $p_x(x_{i+1})$  is computed, and  $H$  is changed according to

$$H_{i+1} = H_i + \gamma_i \frac{s_i s_i^T}{s_i^T y_i} - \frac{H_i y_i y_i^T H_i}{y_i^T H_i y_i} \quad (9)$$

where

$$y_i \equiv p_x(x_{i+1}) - p_x(x_i) \quad (10)$$

The iterative process is then repeated until the minimum is reached.

The DFP method has very fast convergence for nonquadratic as well as quadratic functions if the gradient vector  $p_x$  and the one-dimensional search of minimizing  $p(x_i + \gamma s_i)$  as a function of  $\gamma$  are computed accurately. The success of the one-dimensional search and the accuracy of the numerical gradient  $p_x$  are dependent on the consistency and accuracy with which  $p(x)$  is computed.

One-dimensional search. - Obtaining the one-dimensional minimum of  $p(x_i + \gamma s_i)$  as a function of  $\gamma$  is a critical process for the DFP method. If the one-dimensional minimum is obtained precisely, then equation (3) will be satisfied, and fast, near quadratic convergence of the DFP method will be realized. However, for complicated functions that are encountered, zero for equation (3) is almost impossible. Numerically, it has been found that, even with some error in computation of the one-dimensional minimum, the DFP method will yield near quadratic convergence. The magnitude of this error will be discussed further in this report. It is important, however, that the one-dimensional minimum be efficiently found within this tolerable error.

In the literature, several approaches are described that require computations of  $dp(\gamma)/d\gamma$ , as well as  $p(\gamma)$ , to locate the minimum of  $p(\gamma)$  (refs. 19 to 21). Computations of  $dp/d\gamma$  require the computation of  $p_x(\gamma)$  because

$$\frac{dp}{d\gamma} = p_x(\gamma)^T s_i \quad (11)$$

The time required for computing  $p_x(\gamma)$  prohibits its use in a one-dimensional search. The one-dimensional search described here uses values of  $p(\gamma)$  only.

The one-dimensional search technique features three options: cubic fit using the slope at the initial point (ref. 10), cubic fit without the slope, and the golden section method<sup>7</sup> (ref. 9).

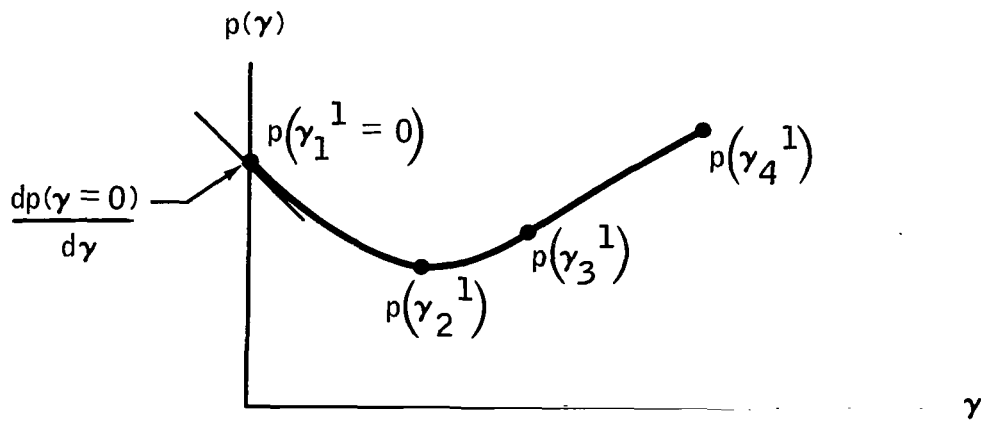
The golden section method is a single-variable, sequential search method used primarily as a starter for the cubic fit method without the slope and as a back-up to both cubic fit methods. First, a search is conducted to find a region of unimodality providing four points:  $[\gamma_j^1, p(\gamma_j^1)]$ ;  $j = 1, 2, 3, 4$  where  $\gamma_1^1 = 0$  and  $[\gamma_2^1, p(\gamma_2^1)]$ ,  $[\gamma_3^1, p(\gamma_3^1)]$ , and  $[\gamma_4^1, p(\gamma_4^1)]$  are such that

$$\left. \begin{aligned} p(\gamma_2^1) &< p(0) \\ p(\gamma_3^1) &< p(\gamma_4^1) \\ \gamma_3^1 &= \gamma_4^1 \frac{\sqrt{5} - 1}{2} \\ \gamma_2^1 &= \gamma_4^1 \frac{3 - \sqrt{5}}{2} \end{aligned} \right\} \quad (12)$$

---

<sup>7</sup>I. L. Johnson and G. E. Meyers, "One-Dimensional Minimization Using Search by Golden Section and Cubic Fit Methods," JSC IN 67-FM-172, Nov. 1967.

An example of this geometrical arrangement is shown in the following diagram.



The golden section method<sup>7</sup> (ref. 9) then proceeds as follows:  $p(\gamma_2^1)$  and  $p(\gamma_3^1)$  are compared. If  $p(\gamma_2^1) < p(\gamma_3^1)$  (as shown), then

$$\left. \begin{aligned} \gamma_4^2 &= \gamma_3^1 & , p(\gamma_4^2) &= p(\gamma_3^1) \\ \gamma_3^2 &= \gamma_2^1 & , p(\gamma_3^2) &= p(\gamma_2^1) \\ \gamma_1^2 &= \gamma_1^1 & , p(\gamma_1^2) &= p(\gamma_1^1) \\ \gamma_2^2 &= \gamma_1^2 + \frac{(\gamma_4^2 - \gamma_1^2)(3 - \sqrt{5})}{2} \end{aligned} \right\} \quad (13)$$

<sup>7</sup>I. L. Johnson and G. E. Meyers, "One-Dimensional Minimization Using Search by Golden Section and Cubic Fit Methods," JSC IN 67-FM-172, Nov. 1967.

and only one additional evaluation of  $p$  is made:  $p(\gamma_2^2)$ . If  $p(\gamma_3^1) < p(\gamma_2^1)$ , then

$$\left. \begin{aligned} \gamma_1^2 &= \gamma_2^1, & p(\gamma_1^2) &= p(\gamma_2^1) \\ \gamma_2^2 &= \gamma_3^1, & p(\gamma_2^2) &= p(\gamma_3^1) \\ \gamma_4^2 &= \gamma_4^1, & p(\gamma_4^2) &= p(\gamma_4^1) \\ \gamma_3^2 &= \gamma_1^2 + \frac{(\gamma_4^2 - \gamma_1^2)(\sqrt{5} - 1)}{2} \end{aligned} \right\} \quad (14)$$

and  $p(\gamma_3^2)$  is calculated. This procedure is repeated until on one iteration (e.g., the  $j$ -th),  $|\gamma_3^j - \gamma_2^j| < \gamma_2^j \cdot \varepsilon$  where  $\varepsilon$  is a prescribed convergence tolerance. The minimum  $\gamma_i$  is computed from

$$\gamma_i = \frac{(\gamma_3^j + \gamma_2^j)}{2} \quad (15)$$

For the cubic fit method without the slope, the four points from equation (12) are retained and are used to begin a search for the minimum of  $p$  using a sequence of third-degree polynomials approximating  $p$ :

$$p(\gamma) \approx a_0 + a_1\gamma + a_2\gamma^2 + a_3\gamma^3 \quad (16)$$

The constants  $a_0$ ,  $a_1$ ,  $a_2$ , and  $a_3$  are obtained by applying the four points to equation (16) and solving the four equations simultaneously for them. From equation (16), the minimums are located from the equation

$$\gamma_j = \frac{-a_1}{a_2 + \sqrt{a_2^2 - 3a_1a_3}} \quad (17)$$

During the search, as a new minimum is found, it is compared to the four points used in the fit; and the point having the largest value of  $p$  among the five is discarded. For successive minimums  $\gamma_j$  and  $\gamma_{j+1}$ , if

$$|\gamma_{j+1} - \gamma_j| < \varepsilon \gamma_j \quad (18)$$

( $\varepsilon$  being a prescribed convergence tolerance), the search is terminated and  $\gamma_i = \gamma_{j+1}$ .

For the cubic fit method using the slope

$$\frac{dp}{d\gamma} (\gamma = 0) \equiv p'(0) \quad (19)$$

in the curve fit, the search procedure begins on the assumption that the function  $p$  is quadratic and that its minimum value is approximated by a lower bound  $p_{LB}$ .

With this information, the first approximation to the minimum (ref. 10) is

$$\gamma_1 = \frac{2[p_{LB} - p(0)]}{p'(0)} \quad (20)$$

The next approximation is also quadratic, using  $p(0)$ ,  $p'(0)$ , and  $p(\gamma_1)$  to calculate the curve fit coefficients. The minimum is approximated by

$$\gamma_2 = \frac{-a_1}{a_2} \quad (21)$$

where  $p(\gamma) \approx a_0 + a_1\gamma + (\frac{1}{2})a_2\gamma^2$ . The next approximation is a cubic fit, and the minimum is found by

$$\gamma_3 = \frac{-2a_1}{a_2 + \sqrt{a_2^2 - 2a_1a_3}} \quad (22)$$

where  $p(\gamma) \approx a_0 + a_1\gamma + (\frac{1}{2})a_2\gamma^2 + (\frac{1}{6})a_3\gamma^3$  and  $a_0$ ,  $a_1$ ,  $a_2$ , and  $a_3$  are found by using the data points  $p(0)$ ,  $p'(0)$ ,  $p(\gamma_1)$ , and  $p(\gamma_2)$ . The previously



generated minimums  $\gamma_2$  and  $\gamma_3$  are compared as in equation (18). If the test in equation (18) is not passed,  $p'(0)$  is dropped as a data point, and  $p(\gamma_3)$  is added; and the search will continue as the cubic fit method without the slope  $p'(0)$ .

If the cubic fit method fails without using  $p'(0)$ , then the search by golden section is used as a backup and initiated with the four points that were retained. Because the golden section method is a much slower converging method than the cubic fit method, it is used primarily as a starter and as a backup.

The speed in finding the first four points in equation (12) is controlled by the selection of  $\gamma_2$  initially. A prescribed constant value is one selection. Another more efficient selection (currently used) is the value of  $\gamma$  at the one-dimensional minimum from the previous iteration  $\gamma_{i-1}$ .

### Penalty Function Method

The exterior penalty function method (refs. 1, 2, 5, and 24) selected for solving the constrained minimization problem is expressed mathematically by

$$p(x) = f(x) + \frac{1}{2} g(x)^T K_1 g(x) + \frac{1}{2} h(x)^T K_2 U(h) h(x) \quad (23)$$

where  $K_1$  and  $K_2$  are diagonal weight matrices, and  $U(h)$  is a diagonal matrix with diagonal elements, unit step functions  $U(h_j)$ .

For positive elements of  $K_1$  and  $K_2$ , equation (23) implies that in minimizing  $p(x)$ ,  $f(x)$  is minimized but is penalized positively for violations in  $g_i$  and  $h_j$ . Increasing each element of  $K_1$  and  $K_2$  causes further restriction on minimizing  $f$ , and more emphasis is placed on decreasing the magnitudes of  $g$  and  $h$ . For increasingly large values of elements of  $K_1$  and  $K_2$ , the solution minimizing equation (23) approaches the solution to the constrained minimization problem (ref. 1). Comparison of the gradient of equation (23) at a minimum of equation (23)

$$p_x = f_x + (g_x, h_x) [K_1 g, K_2 U(h) h]^T = 0 \quad (24)$$

and the equation governing the multipliers for the constrained solution

$$f_x + (g_x, h_x) (\lambda_1, \lambda_2)^T = 0 \quad (25)$$

where  $\lambda$  is a Lagrange multiplier, implies that for increasingly large values of elements of  $K_1$  and  $K_2$  (ref. 1),

$$g_i \rightarrow \frac{\lambda_{1i}}{K_{1i}} \quad (26)$$

and

$$h_j \rightarrow \frac{\lambda_{2j}}{K_{2j}} \quad (27)$$

for the  $h_j$  that are tending to be satisfied by strict equality.

In applications, the size of the elements of  $K_1$  and  $K_2$  must be exploited with regard to computer computation limits. Also, each constraint is identified, and an acceptable violation, which will identify a threshold value for its element in  $K_1$  or  $K_2$ , is assigned to it. By assigning threshold values for the elements of  $K_1$  and  $K_2$ ,  $p(x)$  is minimized, and an acceptable, practical solution to the constrained minimization problem is obtained. For example,  $\bar{v}_f \neq 0$  is a desired value for final

speed. If the constraint for it is defined as  $g = \frac{v_f}{\bar{v}_f} - 1$ , and it is desired that

$v_f - \bar{v}_f \approx \epsilon \bar{v}_f$ , then a reasonable value for its  $K_1$  diagonal element has been found to be  $\approx 10^{-\log_{10} \epsilon}$ .

The penalty function of equation (23) is continuous, together with its first derivatives, if  $f$ ,  $g$ , and  $h$  are continuous; but its second derivatives are discontinuous for all  $x$  such that  $h_j(x) = 0$ . Because  $p(x)$  contains violations in  $h$  and  $g$  at its minimum, an  $h_j$  that is tending toward zero will eventually do so from the positive side during the latter iterations of convergence, and continuous second derivatives will be created during this period.

By using a method requiring gradients to minimize  $p(x)$ , either the closed-form computation approach or a numerical computation approach can be used. If a closed-form approach is taken, then from equation (23),  $p_x$  is computed by

$$p_x = f_x + (g_x, h_x) \begin{bmatrix} K_1 g, K_2 U(h) h \end{bmatrix}^T \quad (28)$$

where  $f_x$ ,  $g_x$ , and  $h_x$  are computed separately and additional computer storage is required. Obtaining  $f_x$ ,  $g_x$ , and  $h_x$  in closed form is often very difficult, if not impossible, and requires storage and many auxiliary computations in solving the nonlinear optimization problems. Also, it is possible that there may be larger errors in computing  $p_x$  than in computing  $p$  because they are not computed from the same equations. Thus, there may be an inconsistency between  $p_x$  and  $p$ . However, consistency between  $p$  and  $p_x$  can be maintained by numerically computing  $p_x$  on the basis of differences of  $p$ . Because only computations of  $p$  are required, no additional storage and auxiliary computations are required. Thus, the penalty function technique becomes completely flexible because any additions or deletions of constraints are made by simply altering  $h$  or  $g$  without having to derive  $h_x$  or  $g_x$ . Several techniques using differences of  $p$  were exploited but proved to be no better than the forward and central difference quotients techniques. For  $\partial p / \partial x_j$ , the forward and central difference techniques are defined, respectively, by

$$\frac{\partial p}{\partial x_j} \approx \frac{p(x_1, \dots, x_j + \Delta x_j, \dots, x_n) - p(x)}{\Delta x_j} \quad (29)$$

and

$$\frac{\partial p}{\partial x_j} \approx \frac{p(x_1, \dots, x_j + \Delta x_j, \dots, x_n) - p(x_1, \dots, x_j - \Delta x_j, \dots, x_n)}{2\Delta x_j} \quad (30)$$

With respect to the Univac 1108 computer, using double-precision arithmetic, non-dimensional state variables, and normalized parameters,

$$0 < |x_j| \leq 10 \quad (31)$$

A suitable choice for  $\Delta x_j$  using forward or central differences is

$$\Delta x_j \approx 1(10)^{-9} |x_j| \quad (32)$$

This perturbation produces partials by using central differences that are accurate to approximately six digits. By using the DFP method, this accuracy was adequate to obtain near quadratic convergence in many instances.

For regions not in a neighborhood of the minimum, the forward difference technique can be used to compute partials accurate enough to produce convergence comparable to that obtained by using central differences, and only half the computations are used in obtaining  $p_x$  (shown in eqs. (29) and (30)). But in a neighborhood of the minimum, central differences are usually required.

## CHARACTERISTICS

To begin the iterative process,  $H_0$  and  $x_0$  are usually chosen by using all the available knowledge and information concerning the minimum. When a general problem such as an ascent problem is solved, information such as the diagonal of a converged  $H$  matrix and a converged parameter vector  $\bar{x}$  is often used to make initial guesses to start related ascent problems. The identity matrix is a popular candidate for  $H_0$ , but many iterations are spent "creeping" toward the minimum while the  $H$  matrix is being manipulated to a point at which significant changes can be made in the parameter vector  $x$ . This is due to the fact that  $H$ , along with  $p_x$ , determines the directions of search (eq. (7)) on each iteration. If  $H$  is completely inappropriate for a given problem, which can happen on the initial guess, the search will not be sufficiently downhill, and short, inefficient searches will result. When  $H$  is the identity matrix, the direction of search is in the local, steepest descent direction. In most cases, this is little consolation since initial local characteristics soon change, and the search is abruptly terminated.

With respect to a Univac 1108 computer, the DFP method has been used successfully in atmospheric flight optimization by applying double-precision arithmetic throughout the search; single-precision arithmetic has been tried, but convergence was very poor (ref. 22). On a seven-parameter problem, a combination of single- and double-precision arithmetic was used with partial success. The author has found that the combination of using single-precision to compute  $p(x)$  and  $p_x$  and double-precision for the DFP method is most successful.

The DFP method does not require the costly matrix inversion computation required by Newton's method and gradient projection. The method is stable if  $p(x)$  and  $p_x$  are computed accurately and consistently and  $H$  is kept positive definite. If  $p(x)$  and  $p_x$  are computed accurately and if one-dimensional searches are accurate,  $H$  will remain positive definite, and the method will maintain stability. Stability, which means that the direction of search  $s_i$  on each iteration is guaranteed to be one in which  $p$  decreases for positive  $\gamma$ , is ensured by the positive definiteness of  $H$ , because

$$\frac{dp(0)}{d\gamma} = -p_x(x_i)^T H_i p_x(x_i) < 0, \text{ for } p_x \neq 0 \quad (33)$$

During the iteration process, it is beneficial to monitor the trend of the convergence. One obvious indicator is the change in the penalty function from iteration to iteration; however, precise trend information cannot be obtained because changes will be made at different, often erratic rates during the normal solution of a problem. Since there is usually no a priori information on the value of the penalty function at the minimum, it is difficult to determine when the solution is found.

The numerical values of the elements of the gradient vector during the iteration also may indicate the efficiency of the iteration. This indication is particularly valuable for ascertaining when acceptable terminal convergence has been achieved because all elements will be zero at the minimum. The primary disadvantage in using this approach is that it requires constant surveillance of a rather large vector in which one element may be the indicator that convergence has not been achieved. Then, too, each element may not decrease monotonically, which makes iteration-to-iteration comparisons difficult. In addition to these difficulties, each element of the gradient vector has its own relative numerical magnitude. A value that may be significantly low for one element may not be sufficiently low for another gradient element.

One of the best indications of the total convergence characteristics of a given problem is the comparison of the directional derivatives  $dp(0)/d\gamma$  at the beginning of each iteration. Another indication of convergence quality is the comparison of  $dp(0)/d\gamma$  and  $dp(\gamma_i)/d\gamma$  for a given iteration.

The directional derivative  $dp(0)/d\gamma$  is a scalar computed from the relationship given in equation (33). It can be interpreted as the derivative of the penalty function in the direction of the one-dimensional search evaluated at the origin of the search. If  $H$  remains positive definite throughout, the DFP formulation ensures that the directional derivative at the origin will increase on each successive iteration (stability). If the derivative is rapidly increasing during initial iterations, it is reasonable to assume that a good estimation was made for  $H$  initially. If the derivative increases slowly for many iterations, this usually indicates that  $H$  is being built. If  $dp/d\gamma$  increases rapidly as the minimum is approached, this indicates that  $p$  is behaving quadratically and near quadratic convergence is possible. If the derivative has increased significantly but then starts to decrease slightly, either the one-dimensional search is not accurate enough or  $p_x$  is inaccurate in this region. A bad one-dimensional search is often characterized by numerous evaluations of  $p$ , possibly indicating a region in which the function is obtaining noisy evaluations, which produce problems with one-dimensional search cutoffs. It could also mean that the one-dimensional search was initiated incorrectly (either too short or too long) and that the search to reach the minimum will be laborious.

During the course of convergence, as the DFP method minimizes  $p$  as expressed by equation (23), the convergence seems to be divided into three regions. The division is indicated by both the directional derivative at the origin and the value of  $p$ . The first region is characterized by large decreases (sometimes two orders of magnitude) in  $p$  on successive iterations. Coinciding with these decreases are rapid changes of  $dp/d\gamma$  ( $\gamma = 0$ ) from one iteration to another. In this region,  $p$  is largely dominated by the last two terms of equation (23). This condition implies that the DFP method is essentially satisfying the point constraints or boundary conditions. The passing of this region of convergence into the second

region of convergence is characterized by much smaller decreases in  $p$  from one iteration to the next, usually in the second or third digit of  $p$ . This result indicates that  $p$  is dominated by  $f(x)$ . In the second region,  $f(x)$  is being minimized, and the boundary conditions are being controlled. Simultaneously,  $dp/d\gamma$  ( $\gamma = 0$ ) changes very slowly and sometimes stays on approximately the same order of magnitude for many iterations. Toward the end of this region,  $p$  will begin to behave quadratically, and stability of the  $H$  matrix will be indicated by each matrix element remaining within approximately an order of magnitude of each element of  $p_{xx}(\bar{x})^{-1}$ . The third region is characterized by rapid increases in  $dp/d\gamma$  ( $\gamma = 0$ ) from one iteration to the next. During this region of convergence,  $p$  behaves like a quadratic function, and on the last iteration, a Newton step will occur with  $\gamma \approx 1$ .

This directional derivative should be zero at the minimum, but no general determination can be made on the smallest acceptable number for complete numerical convergence. This information is not useless, however, because it can be effectively used to give information on the total problem convergence characteristics.

Another characteristic that measures the quality of an individual iteration is the directional derivative  $dp(\gamma_i)/d\gamma$ , the derivative of the penalty function in the direction of the search evaluated after the completion of the search. This directional derivative is particularly useful when compared with its value at the origin  $dp(0)/d\gamma$ . If a one-dimensional search can be assumed accurate but is indicated inaccurate by looking at  $dp(\gamma_i)/d\gamma$  (theoretically should be zero), then  $p_x(\gamma_i) = p_x(x_{i+1})$  is inaccurate because

$$\frac{dp(\gamma_i)}{d\gamma} = s_i^T p_x(x_{i+1}) \quad (34)$$

However, if

$$\left| \frac{dp(\gamma_i)}{d\gamma} \right| < 10^{-4} \left| \frac{dp(0)}{d\gamma} \right| \quad (35)$$

on each iteration, this is usually a good indication that the one-dimensional search and  $p_x$  are being computed accurately enough to ensure near quadratic convergence. If this accuracy is not maintained, the iterative process may eventually degenerate before the minimum is reached. One temporary remedy, which may move the solution a little closer to the minimum, is a series of restarts of  $H$ , because  $H$  has probably gone near singular or indefinite from an accumulation of inaccuracies. But if equation (35) is maintained, a restart of  $H$  is not necessary. In fact, a restart would cause further delay in reaching the minimum because  $H$  would have to reconverge to values near those at its point of restart. To avoid a restart of  $H$  when equation (35) is not satisfied,  $H$  is not updated according to equation (9). This remedy is currently in the program.

As convergence nears the minimum  $\gamma_1 \rightarrow 1$ , this result is seen if equations (5), (7), and (8) are compared. This can be used as a terminal convergence indicator.

## CONCLUDING REMARKS

The Davidon-Fletcher-Powell penalty function method that has been used to solve many problems has proved to be one of the most efficient methods for solving constrained minimization problems. Constrained minimization problems include many different problem definitions concerning parameter optimization, optimal control, boundary value, as well as iterator problems and many others. The Davidon-Fletcher-Powell method is a technique for iteratively solving unconstrained minimization problems; and, for this reason, it is sometimes referred to as an iterator.

To solve problems with continuous state and control inequality constraints, the controls should be modeled parametrically, and the constraints should be handled either explicitly (whenever it is possible to maintain consistency) or by using an error function<sup>6</sup> or a similar method. These types of problems are usually defined through the calculus of variations. However, techniques for solving constrained minimization problems within that realm are usually so complex and inefficient that solutions are often unobtainable.

The primary problems solved with the Davidon-Fletcher-Powell penalty function method have been atmospheric and exoatmospheric flight optimization problems. In solving these problems, many methods were developed for defining and solving optimization problems by parametric methods using the Davidon-Fletcher-Powell penalty function technique.

Lyndon B. Johnson Space Center  
National Aeronautics and Space Administration  
Houston, Texas, March 1, 1976  
986-16-00-00-72

---

<sup>6</sup>I. L. Johnson, "The Error Function Method for Handling Inequality Constraints," JSC IN 73-FM-114, July 1973.

## REFERENCES

1. Kelley, H. J.: Method of Gradients. Ch. 6 of Optimization Techniques, G. Leitmann, ed., Academic Press, 1962.
2. Johnson, I. L.: Impulsive Orbit Transfer Optimization by an Accelerated Gradient Method. J. Spacecraft & Rockets, vol. 6, no. 5, May 1969, pp. 630-632.
3. Schy, A. A.; Adams, W. M., Jr.; and Johnson, K. G.: Computer-Aided Design of Control Systems to Meet Many Requirements. Advances in Control Systems, Conference Proceedings no. 137, North Atlantic Treaty Organization (Neuilly Sur Seine, France), 1973.
4. Johnson, I. L.; and Kamm, J. L.: Parameter Optimization and the Space Shuttle. Paper 8-A-25, Joint Automatic Control Conference (St. Louis, Mo.), Aug. 1971.
5. Johnson, I. L.; and Kamm, J. L.: Near Optimal Shuttle Trajectories Using Accelerated Gradient Methods. Paper 71-328, AAS/AIAA Astrodynamics Specialist Conference (Ft. Lauderdale, Fla.), Aug. 17-19, 1971.
6. Kamm, J. L.; and Johnson, I. L.: Constrained Optimal Ascent-Flyback Shuttle Trajectories. Proceedings of the Fifth International Federation of Automatic Control World Congress (Paris, France), June 12-17, 1972, no. 17.1.
7. Kamm, J. L.; and Johnson, I. L.: Optimal Shuttle Trajectory-Vehicle Design Using Parameter Optimization. Paper 71-329, AAS/AIAA Astrodynamics Specialist Conference (Ft. Lauderdale, Fla.), Aug. 17-19, 1971.
8. Johnson, I. L.: Optimal Rocket Thrust Profile Shaping Using Third-Degree Spline Function Interpolation. Paper 74-823, AIAA Mechanics and Control of Flight Conference (Anaheim, Calif.), Aug. 1974.
9. Wilde, D. J.: Optimum Seeking Methods. Prentice Hall, Inc., 1964.
10. Johnson, I. L., Jr.; and Kamm, J. L.: Accelerating One-Dimensional Searches. AIAA, vol. II, no. 5, May 1973, pp. 757-759.
11. Rosen, J. B.: The Gradient Projection Method for Nonlinear Programming, Part I - Linear Constraints. SIAM J., vol. 8, 1960, pp. 181-217.
12. Rosen, J. B.: The Gradient Projection Method for Nonlinear Programming, Part II - Nonlinear Constraints. SIAM J., vol. 9, 1961, pp. 514-532.
13. Kelley, H. J.; and Speyer, J. L.: Accelerated Gradient Projection. Proceedings of the Symposium on Optimization (Nice, France), June-July 1969, pp. 151-158.



14. Fiacco, A. V.; and McCormick, G. P.: Programming Under Nonlinear Constraints by Unconstrained Minimization: A Primal-Dual Method. RAC-TP-96, Research Analysis Corp., Sept. 1963.
15. Fiacco, A. V.; and McCormick, G. P.: Computational Algorithm for the Sequential Unconstrained Minimization Technique for Nonlinear Programming. Mgt. Sci., vol. 10, no. 4, 1964, pp. 601-617.
16. Brusch, R. G.: A Nonlinear Programming Approach to Space Shuttle Trajectory Optimization. Astronautical Research: Proceedings of the 22nd Congress of the International Astronautical Federation (Brussels, Belgium), Sept. 1971, pp. 185-199.
17. Courant, R.: Variational Methods for the Solution of Problems of Equilibrium and Vibrations. Bull. American Math. Soc., Jan. 1943.
18. Butler, T.; and Martin, A. V.: On a Method of Courant for Minimizing Functionals. J. Math. & Phys., vol. 41, 1962, pp. 291-299.
19. Fletcher, R.; and Reeves, C. M.: Function Minimization by Conjugate Gradients. Computer J., vol. 7, 1964, pp. 149-154.
20. Davidon, W. C.: Variable Metric Method for Minimization. ANL-5990 (rev.), Argonne National Laboratory, Nov. 1959.
21. Fletcher, R.; and Powell, M. J. D.: A Rapidly Convergent Descent Method for Minimization. Computer J., vol. 6, Apr. 1963-Jan. 1964, pp. 163-168.
22. Kelley, H. J.; and Meyers, G. E.: Conjugate Direction Methods for Parameter Optimization. Astronaut. Acta, vol. 16, no. 1, 1971, pp. 45-51.
23. Stewart, G. W., III: A Modification of Davidon's Minimization Method to Accept Difference Approximations of Derivatives. J. Assoc. Computing Machinery, vol. 14, no. 1, 1967, pp. 72-83.
24. Kelley, H. J.; Denham, W. F.; Johnson, I. L.; and Wheatley, P. O.: An Accelerated Gradient Method for Parameter Optimization with Non-Linear Constraints. J. Astronaut. Sci., vol. XIII, no. 4, 1966, pp. 166-169.



510 001 C1 U G 760430 S00903DS  
DEPT OF THE AIR FORCE  
AF WEAPONS LABORATORY  
ATTN: TECHNICAL LIBRARY (SUL)  
KIRTLAND AFB NM 87117

POSTMASTER: 'Undeliverable (Section 158  
Postal Manual) Do Not Return

*"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."*

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

## NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

**TECHNICAL REPORTS:** Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

**TECHNICAL NOTES:** Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

**TECHNICAL MEMORANDUMS:** Information receiving limited distribution because of preliminary data, security classification, or other reasons. Also includes conference proceedings with either limited or unlimited distribution.

**CONTRACTOR REPORTS:** Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

**TECHNICAL TRANSLATIONS:** Information published in a foreign language considered to merit NASA distribution in English.

**SPECIAL PUBLICATIONS:** Information derived from or of value to NASA activities. Publications include final reports of major projects, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

**TECHNOLOGY UTILIZATION PUBLICATIONS:** Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Technology Surveys.

*Details on the availability of these publications may be obtained from:*

**SCIENTIFIC AND TECHNICAL INFORMATION OFFICE**

**NATIONAL AERONAUTICS AND SPACE ADMINISTRATION**

**Washington, D.C. 20546**